Learning by Conceptual Modeling—Changes in Knowledge Structure and Content

Andreas Zitek, Michaela Poppe, Michael Stelzhammer, Susanne Muhar, and Bert Bredeweg

Abstract—The DynaLearn interactive learning environment enables learning by having learners create conceptual models of system’s behavior. This paper reports on exploratory evaluation studies using the DynaLearn software, carried out with learners studying environmental science. Two three-day modeling sessions were conducted in two consecutive years with two students exploring the evolving prototype of the software. The learners worked on assignments designed to achieve specific learning goals. To investigate conceptual changes on behalf of the learner, a set of parameters was applied for semantic text analysis of written pre- and posttests. The evaluation results show key changes occurring in knowledge structure and content in both years for both students. Indications of an effect of prior knowledge on the magnitude of conceptual change were found. The results confirm the potential of DynaLearn for inducing causal and interlinked understanding of environmental systems.

Index Terms—Evaluation study, conceptual knowledge change, learning by modeling, science education

1 INTRODUCTION

Modeling is considered a meaningful and engaging form of learning. The process of creating and testing models of phenomena has been identified as highly relevant for science education [1], [2], [3]. However, the required tools are only sparsely available and often too complex to use, and therefore seldom part of prescribed science education curricula [4]. Particularly, there is a need for software that supports learners in actively dealing with the theoretical concepts involved, e.g., by using visualization and diagrammatic techniques, and by having learners create models and perform concept prediction and explanation [5], [6], [7].

Recently, tools have been developed that allow learners to create such conceptual models [8], [9], [10], based on qualitative reasoning technology [11]. Such tools provide a rich vocabulary for expressing knowledge, particularly concerning cause-effect relationships. They also allow for simulation, typically producing a state-graph representing qualitively distinct behaviors the system manifests.

The DynaLearn interactive learning environment (ILE) is an instance of such a tool [12]. It engages learners in conceptual modeling to develop their understanding of how systems behave and why. A key instrument in DynaLearn is the notion of learning spaces (LSs) [13], a modeling scaffold consisting of a sequence of workspaces with increasing complexity that facilitates a stepwise approach toward developing a system’s view (system’s thinking [14]), while at the same time acquiring the necessary domain-specific knowledge.

To assess the impact of the DynaLearn ILE on acquiring conceptual knowledge, an exploratory evaluation was performed. Two students worked individually with the software guided by assignments. They worked for three days in a row, and did so for two consecutive years, as part of their regular educational activities. For of the evaluation study, instruments were developed to measure the learning progress, and to observe how the learner’s knowledge structure and content changed while working with DynaLearn [15]. The following Hypotheses (H) were investigated:

1. Working with DynaLearn using well-designed assignments will have a significant effect on the knowledge and content structure and its change on behalf of the learners (H1).
2. The vocabulary employed by learners will become more scientific, while progressing to more advanced workspaces in DynaLearn (H2).
3. Prior knowledge will be integrated during learning with new knowledge and revised as needed (H3).
4. Changes in knowledge structure and content can be attributed to the activities performed with the DynaLearn ILE (H4).

The organization of the paper is as follows: First, the DynaLearn ILE is introduced. Then, the evaluation settings and the methods including a description of the different parameters applied to semantic text analysis of pre- and posttests are described. This is followed by a description of the assignments and the results section. Finally, the hypotheses are discussed individually and conclusions are drawn.

2 DYNALEARN ILE

DynaLearn [12], [16] is built on Garp3 [10] and implements a comprehensive vocabulary for learners to express
conceptual notions relevant to reasoning about system’s behavior. Key ingredients to build models in DynaLearn are entities and quantities (see also Fig. 1). Entities represent physical objects or abstract concepts that constitute a system. Configurations represent links between entities. They are used to further express the physical structure of a system. Quantities are associated with entities and describe the dynamics of a system. They typically hold qualitative information concerning the current magnitude and direction of change (derivative), using an interval scale, consisting of an ordered set of labels (without any numerical information). Such a set of labels is called a quantity space.

DynaLearn models capture an explicit notion of causality using the concept of direct and indirect influence [17]. A direct influence can either be positive or negative and represents the initial cause of change. An indirect influence (also known as proportionality) affects its associated quantities according to its direction of change. Hence, the key concept to understand is that influences initiate change while proportionality propagates changes. Additional elements such as correspondences and in/equality statements help to further specify relations between quantities.

Simulations allow for an exploration of the logical consequences of the relationships defined on the quantities in a model. Three important representations of simulation results are: 1) the state graph, showing the possible distinct states of behavior of the simulated system; 2) the value history, the magnitudes and direction of change for all quantities in each state; and 3) the dependency graph, the quantities and their relationships as they apply in each state of behavior in the state graph [18].

DynaLearn provides a number of additional instruments to further support learners. Two of those are relevant for the work presented in this paper: the LSs [13] and the recommendations [19]. The LSs act as a scaffold supporting learners in developing their conceptual knowledge. Each LS provides a restricted set of representational primitives to express knowledge, which focuses, and as such guides, the learner’s knowledge construction process. Moreover, the logical consequences of an expression derived upon simulating provide learners a reflective instrument for evaluating the status of their understanding, to which they can react accordingly. DynaLearn has six LSs, the key characteristics of these are

- **LS1.** Simple concept map (only nodes and links).
- **LS2.** Simple causal model with emphasis on a model structure, entities versus quantities, and the sign of the relationships (positive or negative) between the quantities.
- **LS3.** Causal model (as in LS2) but augmented with the notion of state graph following landmark values characterizing the distinct states of system’s behavior.
- **LS4.** Causal differentiation where the emphasis is on the difference between direct and indirect
influences, and other details regarding the causal explanation underlying the system’s behavior.

- **LS5.** Conditional knowledge where the emphasis is on representing consequences that may only occur under certain conditions.
- **LS6.** Generic and reusable models and model fragments utilizing the full qualitative modeling approach as available in Garp3, including the idea of compositional modeling, and the use of type hierarchy and inheritance.

The second instrument relevant in the context of this paper is the Recommendation: The possibility to obtain feedback on the details captured in the model created by a learner [19]. Recommendation entails two activities: grounding and feedback. DynaLearn guides learners toward proper use of domain specific vocabulary using the grounding step, which refers to the process of linking terms in a learner’s model to concepts in a common vocabulary (DynaLearn uses DBPedia). The learners are in control and decide upon which of the automatically suggested meanings to assign to each of the terms in their models. This is a normative step during which learners deliberately decide upon the meaning they assign to the terms used in their models. The grounding steers the feedback that can be obtained, because the feedback originates from models (stored in an online repository) that have the most similar meaning assigned to their terms. These recommendations represent an individualized feedback to the learner’s model analyzing and presenting differences in organization and content.

There are some noteworthy issues that distinguish DynaLearn from other educational modeling tools. Compared to numerical-based packages such as Coach [20], Insight Maker [21], Causality Lab [22], or even MatLab [23], DynaLearn is fully qualitative. This means that learners actually focus on manipulating conceptual notions and that learners do not have to worry about any numerical detail. The DynaLearn LS1 workspace relates to general concept mapping tools, such as CmapTools [24], while the DynaLearn LS2 workspace relates to tools such as Causal Mapping [25] and Betty’s Brain [26]. However, the representations used within the latter two approaches are less explicit. For instance, these tools do not explicitly distinguish between structure and behavior. Finally, Vmodel [27] relates to the DynaLearn LS4 workspace. A key difference is that Vmodel only allows single-state simulations, while DynaLearn LS4 can also generate a full state graph detailing all possible behaviors of a system. An overall added value of DynaLearn is that it integrates all these different ideas on conceptual modeling into a unified framework for acquiring conceptual knowledge (and accompanying workbench) that facilitates a stepwise approach toward having learners developing their conceptual understanding of how systems behave and why.

## 3 Setting and Method

Two evaluation studies were conducted at the technical secondary high school in Bad Radkersburg (Austria) in April 2010 and March 2011. Each study took three days (from 7:50-13:40 each day, including breaks), and consisted of modeling activities during which LS1, LS2, LS4, and LS5 together with other features of the software were applied. The same students participated in both studies, one female (student 1) and one male (student 2), both 17 years old in 2010. During the studies, the subject matter focused on different types of power production and the effects of those on environment and humans. The first study (2010) focused on wind power plants and the accompanying increase of pump-storage hydropower plants with their effects on the aquatic environment. Due to their prior education, the learners were already informed about power production from renewable resources, but they were not knowledgeable about the potential negative effects of these technologies on rivers and fish. The second study (2011) focused on nuclear power plants and their effects on environment and humans. The subject was triggered by the accident in the nuclear power plant of Fukushima.

At the time of the studies, DynaLearn was still being developed and some unwanted interference with the learning activities due to problems with the software was expected. The students were prepared for this situation by informing them that they were beta testers.

### 3.1 Data Collected

The following data were collected during the studies:

- Subject matter pre- and posttests to investigate the change in the acquired knowledge (free text answering a set of open questions).
- Motivation questionnaires to collect attitudes, impressions, and ideas.
- Full video recordings of the students working on the computer screen, their social interaction, and their questions and answers (from the teacher).
- Complementary information gained by personal note taking during the event, and from the written answers provided by the students.

The results presented in this paper are largely based on data from the pre- and posttests. The other data are used as complementary information.

### 3.2 Transcription and Coding of Tests

For the analysis of the pre- and posttests, quantitative methods were used [28] to characterize changes in knowledge content and knowledge structure [29], [30]. The approach consisted of coding and counting the elements and relationships as used by the learners, to characterize a phenomenon before and after learning activities. Mental maps, extracted from the pre- and posttests, provided additional information on how the knowledge structure changed due to the learning activity [30], [31], [32], [33].

Following these principles, the knowledge structure in pre- and posttests was assessed as a semantic network based on the textual answers provided by students reflecting their understanding [34]. As a prerequisite for analyzing, the test sheets were scanned and made available for Atlas.ti [35]. As a first step, quotations were created in these files capturing keywords (concepts) and causal relations (verbal causal relations, graphical causal relations, wrong causal relations considering both graphical and verbal expressions). Two persons performed the coding
independently. In the case of different judgments, a third opinion was obtained. The codes were exported to a spreadsheet, which provided the first information on conceptual change. Using the coded concepts and causal relations, Atlas.ti was next used to construct generic models of the students’ understanding reflecting their basic knowledge structure in more detail across all written answers in pre- and posttests. When appropriate, central concepts mentioned in the questions were included in these models, when not explicitly mentioned by the student.

### 3.3 Knowledge Content and Structure Analysis

Subsequently, a more detailed classification scheme for the semantic information based on Kurtz dos Santos et al. [36] and Plate [37] was developed. Concepts were divided into entities, attributes, quantities, and processes (landmarks were scored as a specific type of quantity), and their occurrences counted. The explicit mentioning of system behavior notions was also counted. Causal links were either treated as associations, wrong causal links, simple causal links, or causal change links (following the approach by Bunge [38] for analyzing the nature of the links presented in [36]). The assignment of verbal expressions in tests to the corresponding parameters was done on the basis of the initial Atlas.ti codings exported to Excel and the printed original test pdfs containing the codings and the full verbal expressions, and was repeated triply. Between the first two scoring runs, many differences in the judgment occurred between the evaluators, which was only rarely the case between the second and third run. This indicated the growing understanding of what elements are available in the written answers, and how to assign them.

To capture changes in knowledge structure in more detail, the web-like causality was assessed according to Plate [37]. Two measures were used. Link density (LD—ratio of the total number of links/nodes) was calculated from the revised data sets. To better distinguish between the different branching patterns, the web-like causality index (WCI—percentage of nodes with more than one cause or effect) was used. The WCI was calculated using the printed versions of the generic models representing the students’ knowledge as produced with Atlas.ti. Descriptions and examples of all parameters can be found in Table 1. Finally, students mentioning of the central concepts and relations defined as learning goals was also noted.

## 4 ASSIGNMENTS (TREATMENT)

### 4.1 Study 2010

In the 2010 study, the subject matter focused on wind power and fish. Learners used version 0.8.2 of DynaLearn software and performed activities using LS1, LS2, and LS4 (one day each). The two main questions for the learners to investigate were: “Does the production of wind energy influence fish populations?” “How are wind energy production, pump-storage hydropower plants, and hydropeaking related to each other?”

#### 4.1.1 First Day

**Assignments.** After a short introduction to DynaLearn, the pretest was administered. Next, the relationship between wind energy production and the construction of pump-storage hydropower plants was highlighted, and an overview on fish ecological problems caused by hydropower plants was given (PowerPoint presentation). The DynaLearn software was installed and the students received digital background literature. For this assignment, they had to produce a LS1 concept map showing the relationships between alternative renewable energy production forms (focus on wind- and hydropower production) and fish ecology.

**Learning activities and outcome.** Both students were able to produce a comprehensive model (Fig. 2). In LS1, students improved their ecological knowledge mainly by using the additional materials provided (slides of introductory talk, publications, newspaper articles provided as pdf files and the web).

#### 4.1.2 Second Day

**Assignments.** After a short introduction on the importance of understanding system dynamics for making sustainable decisions, both students had to develop a LS2 model on energy consumption and renewable energy production, and potential effects on the environment, with a focus on rivers.

**Learning activities and outcome.** Both students developed a very personal viewpoint on the problem in their LS2 model (Fig. 3). LS2 is well suited for allowing individual viewpoints. Students easily constructed models representing their ideas. This turned out to be very engaging for the students, and formed an interesting basis for additional discussions. During these activities, students often switched back to LS1, to grab the ideas and thoughts from the concept map.

#### 4.1.3 Third Day

**Assignments.** An LS4 example model (generic population) was introduced. As a first modeling assignment, students were asked to rebuild this simple population model, which was for that purpose shown on the classroom presentation device. Based on this generic pattern showing the effect of competing processes on state variables, the possibility of inequality statements and feedback loops was introduced. For the second assignment, the students were asked to build an LS4 model of a control circuit, with a feedback loop, representing the main process (mains frequency regulation) in the wind power production—hydropeaking relationship.

**Learning activities and outcome.** Both students were able to construct an LS4 model with feedback loops representing the basic principle of how mains frequency fluctuations caused by fluctuating wind power production is regulated by hydropeaking.

### 4.2 Study 2011

In the 2011 study, the subject matter focused on nuclear power. Learners used version 0.8.8 of DynaLearn software and performed activities using LS1, LS2, LS4, and LS5 (divided over three days). The main questions for the learners to investigate were: “How does nuclear radiation (in case of an accident in a nuclear power plant) affect the environment and humans?” “How can potential negative effects of radioactive iodine uptake in the thyroid of humans be avoided?” “What is the limit of radioactive
radiation uptake by humans before negative effects are expected?" “What is the structure of a pressurized water reactor and how is energy produced in it?” “What are the ecological effects when cooling water is released to river?”

4.2.1 First Day
Assignments. The first part of the assignment asked students to create an LS1 model. This activity was guided by a short descriptive text on the effects of a nuclear accident on environment and humans, and some additional guiding questions. Students were allowed to use the Internet for additional information. The second part of the assignment asked students to create an LS2 model, based on their LS1 expression.

Learning activities and outcome. Both students were able to build comprehensive LS1 and LS2 (Fig. 4) models. They were also able to identify LS2 as the appropriate LS to implement the textual description.

4.2.2 Second Day
Assignments (2, 3, and 4). At the start of assignment 2, the students received recommendations (derived from an expert model stored in the repository) on their LS2 model built during the first day. The students also had to ground three selected terms using the grounding options provided by DynaLearn. Next, the students had to upload their own model to the repository, obtain recommendations from each other’s model, and again improve their own models.

For assignment 3, the students had to interpret simulation results of a LS4 model dealing with uptake of radioactive iodine by humans, and the potential effect of the intake of iodine tablets. Students had to download the model from the repository, run a simulation, and decide under which circumstances the desirable end situation, a saturation of the thyroid with nonradioactive iodine, is achieved. The main issue was, to run a simulation, identify
the end-states, and compare different simulation pathways, to find out that only the timely uptake of the iodine pill leads to the targeted result.

For the final assignment on this day (assignment 4), the students had to download an incomplete LS5 model from the repository, and to find out why the model was not working. In this faulty model, a conditional fragment was missing. The students had to identify the critical benchmark value for the maximum yearly uptake of nuclear radiation by humans, which should be used to set the landmark for negative effect to occur (as detailed in the conditional expression). The students were told to use the introductory slides containing a complete conditional expression in LS5 of the same model as additional support material.

**Learning activities and outcome.** Following the recommendations, the learners augmented their LS2 models with significant details related to the uptake of radiogenic emissions in different regions of the human body (bones, gastrointestinal tract) and the different types of radiogenic emissions (e.g., plutonium).

The feedback from each others’ models yielded a very dynamic and positive reaction of the students and they got very involved in discussing their model terminologies and their mutual differences in this respect. This session had to be interrupted at a certain point, as students enjoyed it very much to discuss each other’s models.

Both students were able to identify the timely uptake of the iodine pill as the central issue in the case of assignment 3. With regard to assignment 4, after looking up the information with regard to the critical value on the web, both students tried to set the conditional statement. Although both students identified the missing influence...
and proportionality, they were unable to adequately model the conditional expression.

4.2.3 Third Day
Assignments. For the final assignment, two types of nuclear power plants (the boiling water reactor and the pressurized water reactor), and the function of control rods for a controlled production of nuclear energy were introduced. The students had to download an incomplete LS4 model from the repository dealing with 1) the controlled production of energy based on nuclear fission, 2) the principles of a pressurized water reactor with its three different loops, and 3) the subsequent ecological effects of the release of cooling water to river systems. They had to identify the missing elements based on the story that was provided to them on paper, and by additional information they had to look up on the web (some relevant weblinks were provided, especially showing the function of a pressurized water reactor and the effects of cooling water on the river water temperature and its associated organisms). They were asked to compare their models with an expert model in the repository. Finally, they had to run a simulation, and answer questions related to the effect of the position of the control rod.

Learning activities and outcome. Students added the missing information (missing were: entity “Steam turbine” with the quantities “Turbine rotation” and “Energy produced,” and the entity “River” with the quantities “Water temperature” and “Negative effects on aquatic organisms”). Student 1 was able to develop a more detailed model than the expert model even before getting recommendation via the repository based on knowledge gained during earlier modeling activities, whereas the other student did not include the quantity “Water temperature” of the entity “River” (that is causally linked to the effects on aquatic organisms) in his first attempt, but included it after having received recommendations. Both students were able to successfully run simulations and answer the assignment questions.

5 RESULTS
5.1 Study 2010
During the first study, the potential effects of wind energy production on fish and its causal background were explored. The knowledge content of both students appeared to be much more enriched in the posttest, mainly related to quantities (+64%), processes (+76%) and landmarks (+250%) (Table 2). The number of entities decreased for student 2 during the posttest, but quantities and processes increased to twice as much compared to student 1. The number of attributes decreased for both students. Student 2 used a system behavior notion only during posttest.

An increase in causal understanding could be documented by the increase of simple causal links (+128%), but also by an increase of detailed causal links (+89%). The number of associations decreased for student 1 and remained stable for student 2. Wrong causal relations only appeared during the pretest. LD increased for both students (+38 and +51%).
and can be attributed to the increase of causal relations. WCI also increased for both students (+17 and +55%), although the increase was more pronounced for student 2.

Elements contributing to the WCI in the pretest of student 1 were energy production, demand for energy, reservoir power station, hydroelectric power plant, and wind power station \((n = 5)\). During the posttest energy consumption, energy production, reservoir power station, storing of energy, hydroelectric power plant, hydropeaking, habitat, sediment flushing, fish and fish mortality \((n = 10)\) contributed to the increase of the WCI.

In the model reflecting the understanding of student 2, demand for energy, energy production, wind power station, hydroelectric power plant, and water \((n = 5)\) occurred as main elements of the WCI. During the posttest, \(n = 10\) elements had more than one incoming or outgoing causal link or association (energy production, demand for energy, wind power station, reservoir power station, storage of energy, hydroelectric power plant, hydropeaking, flow velocity, habitat, and fish).

For both students, the WCI pointed out that the interaction between the different types of power plants, and the potential negative effect of the related hydropower production mode on habitats and fish was much better understood. The most obvious changes were increasing numbers of simple causal links, processes, quantities, and landmarks between the pre- and posttest. For both students, all the main concepts and relations defined as learning goals appeared in the posttest. Only student 2 explicitly mentioned the regulation of mains frequency as the major goal of the pump-storage power plants, in combination with wind power stations. Student 1 only stated generally that pump-storage power plants are used to compensate for the irregular and unpredictable wind energy production making use of surplus energy produced to pump water to spatially higher positions and to release it when needed. This concept was introduced as the LS4 model focusing on a control circuit as major pattern describing the regulation of the mains frequency.

### 5.2 Study 2011

During the second study, student 1 showed an abundant increase in entities (+42%), whereas the number of entities remained fairly stable for student 2 (Table 3). Student 2 showed a higher increase in quantities (+25%/+18%) and processes (+13%/+0%) compared to student 1, accompanied by an increase in landmarks (+200%/+0%). The number for attributes was low for both students and remained generally stable. Student 1 used two notions of system behavior only in the posttest.

The most abundant changes from pre- to posttest were found for number of causal change links for student 1 (+500%); this parameter did not change for student 2, where the number of simple causal links slightly increased. Wrong causal notions only occurred during the pretest of student 2.

The LD changed in a very similar way in both students (+22%/+29%), whereas the WCI only showed a significant increase for student 2. Student 1 showed a slight decrease in the WCI (−12%/+189%). Elements contributing to the WCI in the pretest of student 1 were nuclear radiation, radioactive element, fuel rod, biological effect on humans, animals and plants, and worker in power plant \((n = 7)\). Elements with more than one incoming or outgoing causal link or association occurring in the posttest were nuclear radiation, nuclear reactor, fuel rod, pressurized water reactor, primary loop, tertiary loop, steam, nuclear meltdown, biological effect on humans, animals and plants and worker in power plant \((n = 10)\). Milli-Sievert as unit to measure the biological effects of nuclear radiation did not show more than one link during the posttest. The changes in central elements for student 1 mainly pointed toward a better understanding of the construction type of a pressurized water reactor.

Elements contributing to the WCI in the pretest were energy, nuclear power plant, nuclear element, and water for student 2 \((n = 4)\); elements contributing to the WCI in the posttest of student 2 were nuclear reactor, radioactive material, core fission, nuclear emission, pressurized water reactor, primary loop, nuclear meltdown, steam, caesium, iodine, strontium, plutonium, cancer, cell, and organs \((n = 15)\). These changes in the knowledge network indicated mainly a change in knowledge related to a generic effect of all of the different radioactive emissions on organs and their ability to cause cancer. In contrast, student 1 linked each specific radioactive isotope only to one specific type of effect (caesium 137 linked to acute radiation sickness, iodine 131 to thyroid cancer and strontium 90 to bone cancer); plutonium was not linked to a specific negative effect on humans. A more detailed understanding of the two different types of nuclear power plants introduced became also evident for student 2. Finally, the scientifically more appropriate term of core fission appeared only in the posttest.

Student 1 did not mention the timely uptake of a iodine pill as potential measure to reduce the uptake of radioactive iodine and, therefore, the risk of getting thyroid cancer (introduced as ready-made LS4 model that needed to be simulated and the results interpreted), whereas student 2 did not mention the effects of nuclear radiation and emissions in case of a nuclear accident on the environment.

### Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre Post %</td>
<td>Pre Post %</td>
<td>Pre Post %</td>
</tr>
<tr>
<td>Entities</td>
<td>24</td>
<td>34</td>
<td>+42</td>
</tr>
<tr>
<td>Quantities</td>
<td>10</td>
<td>13</td>
<td>+18</td>
</tr>
<tr>
<td>Processes</td>
<td>15</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Landmarks</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Attributes</td>
<td>2</td>
<td>3</td>
<td>+50</td>
</tr>
<tr>
<td>S. behaviour</td>
<td>0</td>
<td>2</td>
<td>+200</td>
</tr>
<tr>
<td>Associations</td>
<td>6</td>
<td>12</td>
<td>+100</td>
</tr>
<tr>
<td>Wrong caus.</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Simple caus.</td>
<td>13</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Causal chan.</td>
<td>1</td>
<td>6</td>
<td>+500</td>
</tr>
<tr>
<td>LD</td>
<td>0.39</td>
<td>0.48</td>
<td>+22</td>
</tr>
<tr>
<td>WCI</td>
<td>24.2</td>
<td>21.3</td>
<td>−12</td>
</tr>
</tbody>
</table>

Elements contributing to the WCI in the pretest of student 1 were energy production, demand for energy, reservoir power station, hydroelectric power plant, and wind power station \((n = 5)\). During the posttest energy consumption, energy production, reservoir power station, storing of energy, hydroelectric power plant, hydropeaking, habitat, sediment flushing, fish and fish mortality \((n = 10)\) contributed to the increase of the WCI.

In the model reflecting the understanding of student 2, demand for energy, energy production, wind power station, hydroelectric power plant, and water \((n = 5)\) occurred as main elements of the WCI. During the posttest, \(n = 10\) elements had more than one incoming or outgoing causal link or association (energy production, demand for energy, wind power station, reservoir power station, storage of energy, hydroelectric power plant, hydropeaking, flow velocity, habitat, and fish).

For both students, the WCI pointed out that the interaction between the different types of power plants, and the potential negative effect of the related hydropower production mode on habitats and fish was much better understood. The most obvious changes were increasing numbers of simple causal links, processes, quantities, and landmarks between the pre- and posttest. For both students, all the main concepts and relations defined as learning goals appeared in the posttest. Only student 2 explicitly mentioned the regulation of mains frequency as the major goal of the pump-storage power plants, in combination with wind power stations. Student 1 only stated generally that pump-storage power plants are used to compensate for the irregular and unpredictable wind energy production making use of surplus energy produced to pump water to spatially higher positions and to release it when needed. This concept was introduced as the LS4 model focusing on a control circuit as major pattern describing the regulation of the mains frequency.

During the second study, student 1 showed an abundant increase in entities (+42%), whereas the number of entities remained fairly stable for student 2 (Table 3). Student 2 showed a higher increase in quantities (+25%/+18%) and processes (+13%/+0%) compared to student 1, accompanied by an increase in landmarks (+200%/+0%). The number for attributes was low for both students and remained generally stable. Student 1 used two notions of system behavior only in the posttest.

The most abundant changes from pre- to posttest were found for number of causal change links for student 1 (+500%); this parameter did not change for student 2, where the number of simple causal links slightly increased. Wrong causal notions only occurred during the pretest of student 2.

The LD changed in a very similar way in both students (+22%/+29%), whereas the WCI only showed a significant increase for student 2. Student 1 showed a slight decrease in the WCI (−12%/+189%). Elements contributing to the WCI in the pretest of student 1 were nuclear radiation, radioactive element, fuel rod, biological effect on humans, animals and plants, and worker in power plant \((n = 7)\). Elements with more than one incoming or outgoing causal link or association occurring in the posttest were nuclear radiation, nuclear reactor, fuel rod, pressurized water reactor, primary loop, tertiary loop, steam, nuclear meltdown, biological effect on humans, animals and plants and worker in power plant \((n = 10)\). Milli-Sievert as unit to measure the biological effects of nuclear radiation did not show more than one link during the posttest. The changes in central elements for student 1 mainly pointed toward a better understanding of the construction type of a pressurized water reactor.

Elements contributing to the WCI in the pretest were energy, nuclear power plant, nuclear element, and water for student 2 \((n = 4)\); elements contributing to the WCI in the posttest of student 2 were nuclear reactor, radioactive material, core fission, nuclear emission, pressurized water reactor, primary loop, nuclear meltdown, steam, caesium, iodine, strontium, plutonium, cancer, cell, and organs \((n = 15)\). These changes in the knowledge network indicated mainly a change in knowledge related to a generic effect of all of the different radioactive emissions on organs and their ability to cause cancer. In contrast, student 1 linked each specific radioactive isotope only to one specific type of effect (caesium 137 linked to acute radiation sickness, iodine 131 to thyroid cancer and strontium 90 to bone cancer); plutonium was not linked to a specific negative effect on humans. A more detailed understanding of the two different types of nuclear power plants introduced became also evident for student 2. Finally, the scientifically more appropriate term of core fission appeared only in the posttest.

Student 1 did not mention the timely uptake of a iodine pill as potential measure to reduce the uptake of radioactive iodine and, therefore, the risk of getting thyroid cancer (introduced as ready-made LS4 model that needed to be simulated and the results interpreted), whereas student 2 did not mention the effects of nuclear radiation and emissions in case of a nuclear accident on the environment.
(generically explored within a self-constructed LS2 model), as well as the function of the control rods (introduced explicitly as LS5 model) as important concepts and relations defined as learning goals and treated within the different learning activities. Plutonium, which was added to the LS2 models by both students as missing concept after having got recommendation from the online repository by an expert model, appeared only in both posttests as potential problematic emission in case of a nuclear accident harmful to humans.

Comparing the results of both years indicates a more pronounced increase in most elements, especially in quantities, processes, and simple causal links in 2010. Causal change links were more abundant in 2010, but the percentage of change was higher in 2011. Changes in LD were also more expressed in 2010. The change in the WCI was always higher for student 2 being highest in 2011, whereas for student 1 the WCI was relatively small in 2010 and negative in 2011; although the absolute number of elements with more than one incoming or outgoing causal link or association changed from 7 to 10, the decrease in WCI indicates the addition of more detail and specification in a linear manner.

6 Discussion

6.1 Knowledge Content and Structure (H1 and H2)
The assessment of the effect of different learning activities on conceptual change supported by the DynaLearn ILE yielded relevant changes in knowledge structure and content. Most of the measures indicated a much better and more detailed understanding of the environmental issues treated within the different learning activities. Especially, the WCI as measure of the number of complex nodes was useful to identify changes related to central concepts in the systems taught. Simple causal links and causal change links appeared to change more significantly in the 2010 study. Generally, conceptual change was expressed more in 2010, when the students’ prior knowledge on the subject matter was limited. Conceptual change was characterized by the refinement of some relations (wrong causal expression), and the refinement of vocabulary. A better process-based understanding was developed. This was more obvious when the viewpoint initial held by the students was more static. The complexity of the models representing their understanding also increased.

Self-oriented learning became evident during different learning activities as both students developed a different focus on the environmental problems caused by power production (2010 study), mainly during their LS2 modeling activity. Student 1 was focusing more on ways to minimize energy consumption, whereas student 2 focused on the degradation of the environment by the increasing need for energy production caused by human energy consumption.

All of the major types of conceptual change that have been described by Driver and Oldham [39] (accretion— the addition of parts to an existing structure, tuning— involving small modifications to an existing structure, and restructuring— involving major changes in the structuring of knowledge) could be documented. In accordance with this are also the two types of conceptual change described by Duit and Treagust [40] (weak knowledge restructuring, assimilation or conceptual capture and strong/radical conceptual knowledge restructuring, accommodation or conceptual exchange).

A deeper analysis of the language specific causal linguistic coding of the written answers according to Wolff [41] and Wolff et al. [42] was out of the scope of this study.

6.2 Integration of Prior Knowledge (H3)
Prior knowledge generally was more linear, which was mainly documented by increases in LD and WCI. However, the initial structure of prior knowledge of a student plays a role for the potential conceptual changes that might occur [43]. For example, in the second year, the WCI decrease indicated a well-integrated and linked initial knowledge structure, where additional elements were mainly added in a linear manner.

However, some relevant elements defined as learning goals (mains frequency as major element in the wind power—pump-storage power plant relationship (2010 study) by student 1 and the effects of nuclear radiation and emissions in case of a nuclear accident on the environment as well as the function of the control rods by student 2 were not mentioned during posttests. It is unclear if these elements were forgotten to be answered during posttest situations (student 2 in both years hurried while completing the posttest), or if this really represented a lack of knowledge. During the modeling activity related to the exploration of mains frequency in LS4, student 1 was very interested in the different ways of energy consumption in relation to the need for adequate energy production, and did focus mainly on the application of the + calculus to add up the different consumptive elements (industry, household, etc.) to relate the outcome to different modes of energy production. It could be that this focus interfered with the desired focus on mains frequency as the defined learning target. Interviews or talk-aloud protocols may help to identify the underlying reasons for this phenomenon.

Prior knowledge was integrated with new information, and only some elements were missing in the posttests. This was the case especially for the topic on nuclear garbage, which appeared in the pretest of one student. This element was never explicitly requested in assignments. This points out the importance of an adequate development of tasks, and also especially with the design of assignment questions, as this may influence the way prior knowledge is taken into account. It also could be that the questions of the tests were formulated to broad, and therefore, some knowledge elements were left out in the posttests. A more specific formulation of test questions might help to better retrieve the appropriate mental model on behalf of the students. Motivational issues might be responsible for producing these effects on the results [44], [45].

It is known that learning can be significantly enhanced when it addresses the students prior and alternative conceptions [46]. Expert teachers are highly aware of the importance to consider prior knowledge in instruction [47]. Teachers therefore not only need to introduce the desired conception but also consider the alternative conceptions of students [48]. DynaLearn explicitly builds on the existing
understanding of an environmental issue and its extension/ restructuring by new information, achieved mainly via the recommendations from an expert model. Grounding supports the transgression from personal vocabulary to a more scientific one. In this sense, DynaLearn is in line with the common understanding that “learning consists of the enrichment and reorganization of existing knowledge structures” [49].

6.3 Changes Attributed to DynaLearn (H4)
Most observed knowledge changes could be attributed to different learning activities. For example, plutonium as potential radioactive emission in the case of a nuclear accident was introduced in both LS2 learner models (2011 study) based on the feedback from an expert model in the repository and only showed up in the posttests. The detailed structural information about the pressurized water reactor (2011 study) was introduced at the third day during assignment 5. This points out the relevance of combining different learning activities taking advantage of the possibilities the DynaLearn ILE offers to achieve specific learning outcomes.

7 Conclusion
DynaLearn offers a systematic framework, which allows the expression of prior knowledge and perceptions, and to relate this systematically to new information. The learner is hereby guided via interactive learning routes from initial personal understanding toward a more scientific understanding. The studies presented in this paper show the general applicability of the DynaLearn ILE to successfully convey conceptual understanding of environmental systems. The set of measures applied was well suited to document the major elements of conceptual change. To better understand the effects of the specific modeling language on the final causal linguistic coding of written answers, a deeper semantic analysis is needed. Together with a detailed analysis of the technological factors making the tool an improvement over previous attempts, this will provide further guidance for the development of new causal learning technologies focused around the natural ability of humans to build up causal understanding of their environment. Further studies including more subjects (>20 students) and a control group are needed to specifically enlighten the benefits of this new technology. The effectiveness of the presented learning approach, especially for self-directed learning and development of higher order reasoning skills (e.g., transfer of patterns), still needs to be further evaluated.

Acknowledgments
The authors thank Dr. Josef Maßwohl, headmaster of the International College on Electrical Engineering with Specialization on Renewable Energies (iHTL) in Bad Radkersburg, Styria, Austria, for his kind collaboration, and Sophie Fastian and Oliver Kindler for their active participation in the evaluations. The research presented here was cofunded by the European Commission within the Seventh Framework Programme for Research and Development (FP7, project DynaLearn, number 231526, http://www.DynaLearn.eu).

References
Andreas Zitek received the MSc degree. He is currently a senior researcher at the University of Natural Resources and Life Sciences (BOKU) Vienna, situated at the new Research and University Center in Tulln. From 1999 through 2012, he was with the Institute of Hydrobiology and Aquatic Ecosystem Management of BOKU. Since 2009, he has been a group leader of the Aquatic Ecosystem Chemistry Group at the VIRIS Laboratory, Department of Environmental Management, University of Natural Resources and Life Sciences, Vienna. His work deals with the assessment and evaluation of different types of riverine landscapes. Another important focus in her work is the application and evaluation of new learning and teaching approaches at the school and university level and the transmission of scientific knowledge to the younger generation.

Michaela Poppe is a senior researcher at the Institute of Hydrobiology and Aquatic Ecosystem Management, University of Natural Resources and Life Sciences, Vienna. Her work focuses mainly on the history of riverine landscapes, the planning of river restorations, depicting riverine functions and processes, and the application and evaluation of new learning and teaching approaches at the school and university level.

Susanne Muhar is an associate professor at the Institute of Hydrobiology and Aquatic Ecosystem Management, University of Natural Resources and Life Sciences, Vienna, where she leads the Riverine Landscape Working Group. The Aquatic Landscapes Group covers a broad range of scientific disciplines (including landscape planning, history, and geography) and closely collaborates with the working groups on fish ecology and benthic river ecology. Her work focuses on the integrative and sustainable management of river catchments as well as the integrative and reflected education of students.

Bert Bredeweg is an associate professor and leader of the Qualitative Reasoning (QR) Group, which is part of the Theory of Computer Science, one of the eight research groups within the Informatics Institute at the University of Amsterdam. The QR Group is widely recognized for their work on qualitative modeling and simulation. He coordinates the International DynaLearn Project.